

## Durham Research Online

---

### Deposited in DRO:

08 December 2020

### Version of attached file:

Accepted Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Medeiros, Andrew S. and Milošević, Đurađ and Francis, Donna R. and Maddison, Eleanor and Woodroffe, Sarah and Long, Antony and Walker, Ian R. and Hamerlík, Ladislav and Quinlan, Roberto and Langdon, Peter and Brodersen, Klaus P. and Axford, Yarrow (2021) 'Arctic chironomids of the northwest North Atlantic reflect environmental and biogeographic gradients.', *Journal of biogeography.*, 48 (3). pp. 511-525.

### Further information on publisher's website:

<https://doi.org/10.1111/jbi.14015>

### Publisher's copyright statement:

This is the peer reviewed version of the following article: Medeiros, Andrew S., Milošević, Đurađ, Francis, Donna R., Maddison, Eleanor, Woodroffe, Sarah, Long, Antony, Walker, Ian R., Hamerlík, Ladislav, Quinlan, Roberto, Langdon, Peter, Brodersen, Klaus P. Axford, Yarrow (2021). Arctic chironomids of the northwest North Atlantic reflect environmental and biogeographic gradients. *Journal of Biogeography* 48(3): 511-525. <https://doi.org/10.1111/jbi.14015>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

### Additional information:

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

# **The distribution of arctic Chironomidae (Insecta: Diptera) in the northwest North Atlantic region follows environmental and biogeographic gradients**

Andrew S. Medeiros<sup>1</sup>, Djuradj Milošević<sup>1,2</sup>, Donna R. Francis<sup>3</sup>, Eleanor Maddison<sup>4</sup>, Sarah Woodroffe<sup>4</sup>, Antony Long<sup>4</sup>, Ian R. Walker<sup>5</sup>, Ladislav Hamerlík<sup>6,7</sup>, Roberto Quinlan<sup>8</sup>, Peter Langdon<sup>9</sup>, Klaus Peter Brodersen<sup>10</sup>, and Yarrow Axford<sup>11</sup>

1. Dalhousie University, School for Resource and Environmental Studies, College of Sustainability, Halifax, Nova Scotia, Canada
2. University of Niš, Department of Biology and Ecology, Niš, Serbia
3. University of Massachusetts, Department of Geoscience, Amherst, Massachusetts, USA
4. Durham University, Department of Geography, Durham, UK
5. University of British Columbia Okanagan, Departments of Biology, and Earth, Environmental and Geographic Sciences, Kelowna, British Columbia, Canada
6. Matej Bel University, Faculty of Natural Sciences, Banská Bystrica, Slovakia
7. Institute of Zoology, Slovak Academy of Sciences, Bratislava, Slovakia
8. York University, Department of Biology, Toronto, Ontario, Canada
9. University of Southampton, School of Geography and Environmental Science, Southampton, UK
10. Raadvad Nature School, Rudersdal Kommune, Denmark
11. Northwestern University, Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL, USA

\* Author for correspondence: Andrew S. Medeiros; [andrew.medeiros@dal.ca](mailto:andrew.medeiros@dal.ca)

**Key words:** Arctic, biogeography, chironomids, freshwater ecosystems, gradient analysis, non-biting midges, Self Organizing Map

**Paper type:** Primary Research Article

## **Abstract**

### *Aim*

While we understand broad climate drivers of insect distributions throughout the Arctic, less is known about the role of spatial processes in determining these relationships. As such, there is a need to understand how spatial controls may influence our interpretations of chironomid environment relationships. Here, we evaluated whether the distribution of chironomids followed spatial gradients, or were primarily controlled by environmental factors.

### *Location*

Eastern Canadian Arctic, Greenland, Iceland.

### *Taxon*

Non-biting midges (Chironomidae)

### *Methods*

We examined chironomid assemblages from 239 lakes in the western North Atlantic Arctic region (specifically from the Arctic Archipelago of Canada, two parts of west Greenland (the southwest and central west), and northwest Iceland). We used a combination of unconstrained ordination (Self Organizing Maps); a simple method with only one data matrix (community data), and constrained ordination (Redundancy Analysis); a canonical ordination with two data sets where we extracted structure of community related to environmental data. These methods allowed us to model chironomid assemblages across a large bioregional dimension and identify specific differences between regions that were defined by common taxa represented across all

regions in high frequencies, as well as rare taxa distinctive to each region found in low frequencies. We then evaluated the relative importance of spatial processes versus local environmental factors.

### *Results*

We find that environmental controls explained the largest amount of variation in chironomid assemblages within each region, and that spatial controls are only significant when crossing between regions. Broad-scale biogeographic effects on chironomid distributions are reflected by the distinct differences between chironomid assemblages of Iceland, central-west Greenland, and eastern Canada, defined by the presence of certain common and low-frequency, rare taxa for each region. Environmental gradients, especially temperature, defined species distributions within each region, whereas spatial processes combine with environmental gradients in determining what mix of species characterizes each broad and geographically distinct island region in our study.

### *Main conclusions*

We outline that while biogeographic context is important for defining interpretations of environmental controls on species distributions, the primary control on distributions within regions is environmental. These influences are fundamentally important to know for reconstructing past environmental change and better understanding historical distributions of these insect indicators.

## Introduction

Of the aquatic insects that occur in the Arctic, the order Diptera dominate in number of species and sheer abundance. Indeed, chironomids (Diptera: Chironomidae) make up between one-fifth and one-half of the total number of arctic insect species (Oliver, 1971). Across the Arctic, the relative abundances of different chironomid subfamilies have been clearly related to temperature (e.g., Walker et al., 1991; Brooks and Birks, 2001; Medeiros and Quinlan, 2011; Eggermont and Heiri, 2011; Dickson and Walker, 2015; Nazarova et al., 2015). Species of the tribe Chironomini, for example, are rare or absent at the highest latitudes, whereas those of the subfamily Orthoclaadiinae are most abundant at the coldest sites and decline in warmer regions (Oliver, 1971; Gajewski et al., 2005; Quinlan et al., 2005; Fortin et al., 2015). Thus, the distribution of chironomid species can be informative in indicating past or ongoing temperature change, and subfossil chironomid assemblages preserved in lake sediments have become a useful tool for reconstructing arctic temperatures of the Holocene and late Pleistocene (e.g., Medeiros et al., 2012; Kaufman et al., 2016; Engels et al., 2020).

The composition of chironomid communities across the circumpolar Arctic also has well-demonstrated dependence on other local environmental factors like nutrient conditions and hypolimnetic oxygen, particularly when encountering extremes of these conditions (Brodersen and Anderson, 2002; Medeiros and Quinlan, 2011; Medeiros et al., 2015; Luoto and Ojala, 2018). Likewise, the role of geological provinces come into play across large spatial gradients (Westover et al., 2009), as seen in research on benthic invertebrates by the absence of Mollusca & Gastropoda in regions with igneous bedrock and calcium-poor waters (Namayandeh and

Quinlan, 2011). Less is known about how aquatic invertebrate species distributions have been affected by biogeographical constraints that likely limit dispersal and/or the ability to colonize new habitats in the Arctic. For example, Taylor et al. (2016) found that the prevalence of recent observations of another midge, *Chaoborus* (Chaoboridae), in tundra regions of the Seward Peninsula, Alaska, were likely due to dispersal from warmer regions across ‘stepping stone’ habitats (small, shallow, and warmer ponds) that had recently developed due to increased thermokarst processes. Identifying potential limits on the distribution of insects across large spatial scales is also necessary for understanding the ongoing changes in species distribution that may result from modern climatic and environmental change (Primack et al., 2018; Engels et al., 2020).

Most of the presently ice-free land in the northwest North Atlantic region – Iceland, Greenland, and the eastern Canadian Arctic – was covered by continental ice sheets during the last glacial maximum, which extirpated local populations of many plants and animals and required dispersal over long distances in the Holocene (Bennike, 1999; Crump et al., 2019). It has long been debated what plants and animals, if any, survived through the last glacial period in these regions (Buckland and Dugmore, 1991; Bergersen, 1995; Bennike, 1999; Rundgren and Ingolfsson, 1999; Brochmann et al., 2003). In addition, due to their island nature and potential dispersal connections to both the Palearctic and Nearctic, these regions are biogeographically very complex (e.g., Bergersen, 1995; Johansen and Hytteborn, 2008; Novichkova et al., 2014; Panagiotakopulu, 2014) and are strong candidates for studying dispersal limitations on chironomid assemblages.

The mechanisms of biogeographic limitations on dispersal for arctic dipterans may be related to regionality in (sometimes extreme) environmental conditions, distance between available habitats, and even the mating process itself. The obvious limitation of distance for dispersal of chironomids is compounded by large oceanic barriers between the continents and islands, gradients in elevation that can span 1000 m, the harsh environment (e.g., winds, discussed below), and salinity of coastal lakes and ponds (Dickson et al., 2014). For example, the chironomid fauna of Ireland is relatively depauperate compared to the European mainland, and dispersal barriers to post-glacial colonization, including distance to Great Britain & the mainland along with a dominant southwesterly wind direction, have played an important role in biodiversity levels and community composition (Murray & O'Connor, 1992). Dispersal and colonization of new habitats often occur during regular movement within the mating process, such as the passive dispersal of fertilized females (Oliver, 1971). The mating swarm is not typically a dispersal phase, but the common high winds in arctic environments can disperse entire populations away from a particular area (Beattie, 1981) and wind is likely an important means of dispersal for dipterans in the North Atlantic region (Panagiotakopulu, 2014). Wind can also hinder the ability of some chironomid species to fly (causing them to land), while assisting normally flight-limited terrestrial species with dispersal (Delettre, 1988). Medeiros and Quinlan (2011) noted several temperate species of the tribe Chironomini on Bylot Island, isolated from other members of that tribe next occurring 500+ km to the south. These types of assemblages could exist due to animal vectors, such as migratory waterfowl that transport propagules to their breeding colonies (Frisch et al., 2007). Wind-mediated transport of chironomid larvae inside moss tufts may provide another mechanism facilitating colonization of new habitats in regions

with high amount of mosses and strong wind activity, such as Arctic and alpine regions (Bitušík et al., 2017). In addition to wind and migratory birds, ice-rafting and driftwood have also been proposed as important vectors for species migration within the North Atlantic Arctic region (Buckland and Dugmore, 1991; Bennike, 1999; Johansen and Hytteborn, 2008). Indeed, analysis of the genetic composition of chironomids from eastern Greenland suggest both a small endemic population as well as origins of palearctic dispersal (Ekrem et al., 2018). Thus, both the limitation of dispersal and the ultimate colonization potential of sites across the Arctic are likely complicated and time-dependent.

In order to assess spatial controls on chironomid distributions in this interesting region, we assemble for the first time a large dataset of modern chironomid assemblages from 239 lakes in the Arctic Archipelago of Canada, southwest and central west Greenland, and northwest Iceland. We model chironomid assemblages across a large bioregional dimension, in order to define indicator taxa representative of each bioregional component, and to assess the relative importance of spatial processes versus local environmental factors in governing chironomid assemblages between regions. Such biogeographic-distance related limits on insect dispersal have implications for chironomid-environment relations across large spatial gradients, and thus for paleoecological models that assume that species assemblages are in equilibrium with climate over time.

## **Materials and methods**

### **Data sources**



We analyzed a large dataset of chironomid assemblages from 239 lakes across the Canadian Arctic Archipelago, southwest and central west Greenland, and Iceland. The sites span from 61.68 to 80.81 °N and -21.26 to -113.91° W, covering a gradient in annual mean temperatures from 4.4 to -21.8 °C and altitude from 1 to 855 m a.s.l. (Figure 1). Much of our study region is remote and not easily sampled, so we utilize pre-existing collections from published studies (except for data from Nuuk, Greenland, which we publish here for the first time). Because assemblages in these samples were previously enumerated by multiple investigators employing varying degrees of taxonomic resolution, samples underwent taxonomic harmonization using Brooks et al. (2007) (unless otherwise noted) as the primary identification metric. Harmonization of taxonomy between datasets was confirmed through re-numeration of samples across datasets and/or collaborative discussion of identifications with the original authors of each dataset. Most studies used a threshold of 50 head capsules for inclusion in their datasets with the exception of Gajewski et al. (2005) who used a threshold of 35 head capsules for high Arctic sites lacking sufficient abundance. Within the 239 lake dataset the median abundance of head capsules in each lake sample was 93.5.

Each regional training set was collected as part of its own separate study, using slightly different field protocols to collect undisturbed surface sediments. In central west Greenland, surface-sediment samples from the deepest part of 42 lakes were collected between 1996 and 2000. The uppermost 0–1 cm sediment layer was taken using a HON-Kajak sediment corer and stored in plastic bags. For further details on the collection of environmental data and sample processing see Brodersen and Anderson (2002).

The surface samples from southwest Greenland were collected in summer 2011 from 25 lakes in the Nuup Kangerlua area, ~80 km east of Nuuk using a Hongve corer (Wright, 1990). Samples were taken from the deepest part of each lake where possible, as determined by echo sounding. Chironomid head capsules were extracted from the surface sediment (0–1 cm) following the protocol detailed in Brooks et al. (2007). pH, conductivity, dissolved oxygen and water temperature were measured at each sample site using a YSI multi parameter sonde at 0.5m water depth and 0.5 m above the lake bed (Supplementary Data). Water samples collected for laboratory analyses were collected at 0.5 m water depth at the same location as the *in situ* measurements. These samples were filtered, kept cool and stored in the dark whilst in the field, and were analysed in the laboratory for alkalinity, total nitrogen and phosphorus, major anions and cations, trace metals, and organic constituents (Clesceri et al., 1998). Filtrate was analysed for chlorophyll-a. Transparency of the lake waters were measured with a Secchi disk. Surface sediments (0–1 cm) were analysed for percentage loss-on-ignition following standard methods (Dean, 1974; Heiri et al., 2001). Although only selected variables of the collected environmental data are used in our analyses, to maintain consistency between training sets, the full suite of environmental parameters are reported in Supplementary Data.

The collection of Icelandic samples and environmental data in 2002 and 2003 is outlined in Langdon et al. (2008). Baffin Island samples were collected during summers of 1994 through 1996. Details of sample collection and processing protocols can be found in Francis et al. (2006), and (referenced therein) Joynt and Wolfe (2001). In order to standardize the environmental data used in our analysis, and because the remoteness of many of the training set sites precludes availability of reliable local meteorological measurements, mean annual air temperature and

mean annual precipitation, and mean annual temperature of warmest quarter were extracted for each lake across all training sets using the WorldClim 2.0 gridded bioclimatic dataset (Fick and Hijmans, 2017). For the Icelandic and southwest Greenland training sets in particular, because the original studies by design sampled along elevation gradients to capture temperature gradients within small but high-relief geographic areas, the use of WorldClim data will introduce some uncertainty due to the co-linear nature of elevation and temperature, as well as the close proximity of sites.

## **Data Analysis**

Chironomid assemblages characterized by enumerated chironomid head capsules were modeled with an unsupervised artificial neural network (ANN) and a Self-Organizing Map (SOM) approach (Kohonen, 1982; Chon et al., 1996). The SOM approach is used for classification and prediction aspects for biomonitoring approaches and presentation of assemblage patterns of chironomids across habitats (Milošević et al., 2013). In comparison to traditional multivariate methods, the SOM can process large data sets, making full use of the available 2d space to model non-linear variability, and is less affected by outliers. In addition, the SOM can visualize the relationship between explanatory (environmental data) and response (community data) variables in a way that maintains an unconstrained ordination of response variables (Park et al., 2004). This visualization technique is composed of three layers; input, output, and hidden layers. To obtain the ordination pattern of sampling sites, the SOM was trained from the input data, where the number of neurons in the input layer was determined by the number of input vectors (number

of sampling sites). The input data was rescaled with a log transformation ( $\log(x + 1)$ ) and each input vector (sampling site) was sent through the network within an unsupervised learning process. As a result of this process each vector was attached to a neuron in the output neural network, with the position of each site determined by chironomid assemblage dissimilarity based upon Euclidean distance. The distance between neurons on the two-dimensional neural network indicates the dissimilarity between chironomid assemblages. In other words, sampling sites with similar chironomid assemblages are located in the same or adjacent neurons, while sampling sites with very different assemblages are placed in distant neurons of the two-dimensional neural network. To classify sampling sites into groups (and detect regionality), the neurons on the trained SOM are distinguished into subsets with similar chironomid assemblages using the *k*-means method (Jain and Dubes, 1988). The classification patterning was followed by a distance-based permutational multivariate analysis of variance (PERMANOVA), in PRIMER v 6 (Clarke and Warwick, 2001), on the Bray–Curtis resemblance matrices, which tested the significant differences in chironomid assemblages between the clustered outputs with the *k*-means method. The test was conducted with the design of one factor, SOM groups (with 5 levels; Fig 2), using 9999 permutations to estimate the *p*-value of pseudo-*F* statistic.

Species sorting of samples was visualized via a component plane, constructed by the SOM algorithm. Each environmental parameter was introduced in the trained SOM passively, not influencing previous ordination and classification processes. An output of the distributional pattern of each parameters' intensity was visualized, enabling the visual correlation with other active entities (ordinated sampling sites and chironomid taxa) of the SOM model. If a clear distributional gradient appeared on the SOM it suggests that the parameter governs the

variability in community structure. However, since SOM is an exploratory method, the variability of environmental parameters between SOM groups was confirmed with a non-parametric ANOVA (Kruskal–Wallis tests). For pair-wise comparisons, the Mann–Whitney post-hoc test with Bonferroni correction for multiple comparisons ( $p < 0.01$ ) was applied.

For the SOM approach, the number of output neurons (map resolution) determines the size of the 2d map space for data ordination and classification. The map resolution is an important parameter that can influence how precise the detection of data deviation can be. If the resolution is too low or too high, important differences in data variability could be too small for valid interpretation or omitted, respectively. Therefore, we chose a  $9 \times 8$  map resolution as the most appropriate context for the data as determined using standard methods (Vesanto et al., 2000; Park et al., 2003) with a low number of empty output neurons (Penczak et al., 2012). The empty neurons are units without assigned samples and their high number in the map can hinder the gradient analysis via component planes. The empty neurons in this study were ignored in further statistical analysis. To define indicator taxa most responsible for classifications of chironomid communities into SOM groups, the IndVal indicator species analysis approach was conducted using PC-ORD 6.08 (McCune & Mefford 2011). The frequency and abundance of all taxa were tested between the SOM groups, where if taxa were present  $>50\%$  of the time, the indicator species had IndVal value  $>25$  and were considered representative (common taxa) of the SOM group. If taxa were present  $< 50\%$  of the time, the IndVal value would be  $< 25$ ; we consider these taxa rare (less frequent). Using the component plane, the distribution of abundance of all representative taxa was then visualized over the trained SOM.

The multiscale distance relationship among sampling sites within each region was modeled and described via spatial variables using the distance-based Moran's eigenvector maps (dbMEMs; Borcard and Legendre, 2002; Borcard et al., 2004; Dray et al., 2006). These variables describe the spatial variation along a spatial scale, where high eigenvalues of dbMEMs indicate broad-scale patterns of relationships between studied sites while dbMEMs with low eigenvalues represent fine-scale relationships (Gronroos et al., 2013). This method uses geographical coordinates and obtains the orthogonal spatial variables as an output. Only dbMEMs with positive autocorrelation were analyzed further for variance partitioning.

The relative importance of spatial processes and local environmental factors in chironomid community structuring was subsequently tested by variance partitioning based on redundancy analysis (RDA) for each SOM group identified using the function 'varpart' in the R vegan package (Oksanen et al., 2007). Prior to RDA, chironomid abundance data were Hellinger transformed (Legendre and Gallagher, 2001). The environmental parameters included in RDA were selected using forward selection with two criteria, exceeding the critical  $p$ -value ( $p=0.05$ ) and exceeding the adjusted  $R^2$  value of the global RDA model (Blanchet et al. 2008) using the function 'R2adj' in the R packfor package (Dray et al. 2009). Adjusted  $R^2$  values are reported presenting the unbiased estimates of explained variation (Peres-Neto et al., 2006). The statistic design was set to (Spatial (S) + Environmental (E) parameters), pure (E|S and S|E) and shared (S and E) effects of explanatory variable groups on the response variables (i.e., the chironomid assemblage). The significance of total and pure effects was tested using the function 'anova' in the R vegan package (Oksanen et al., 2007).

The relative power of analysis strongly depends on the similarity/dissimilarity measure used. Despite the double-zero problem in species abundance data (Legendre and Legendre 1998; Legendre and Gallagher, 2001), which constrains using Euclidean distance in linear methods (e.g. PCA and RDA), the SOM approach deploys non-linear transformation to address distributions that are skewed due to the absence of taxa. This approach has been demonstrated in studies that have examined distributions of algae (Joo and Jeong, 2005), macroinvertebrates (Chon et al., 1996; Milošević et al. 2013; Park et al., 2006) and fish (Brosse et al., 2001; Kruk, 2007; Stojković et al., 2013). However, for the conventional linear approaches used in this study, we also used pre-transformation (Hellinger transformation) and dissimilarity measures (Bray–Curtis distance) for RDA and PERMANOVA methods, respectively, in order to address the double-zero effects that are common influences for Arctic regions due to low diversity.

## **Results**

The SOM classified our 239 sample sites into five groups (Fig 2). Each SOM group was characterized by distinctive assemblage structure, where structural changes between all five groups were significant (Pseudo- $F=40.149$ ,  $p=0.0001$ ). The classification mainly grouped sites by region, with Greenland and Iceland each representing a SOM cluster, and the Canadian Archipelago forming 3 distinct groups.

IndVal analysis found 41 taxa with IndVal  $>25$ , identifying taxa that can be considered common and representative for each of the SOM groups (Fig3a, Table 1). In addition, a further 36 taxa in the model had IndVal  $<25$  ( $p < 0.05$ , Table 1), indicating that they appeared significantly more often in one of the SOM groups but with low relative frequency (considered

important but rare taxa). A Kruskal–Wallis test confirmed a significant difference in environmental parameters between SOM groups (Table 2). Likewise, passively introduced environmental variables indicated a clear gradient on the component planes, which were concordant to the classification pattern generated by the SOM exclusively from chironomid assemblages (Fig 3b). This was especially true for mean and summer air temperatures (temperature of the warmest quarter), with groups IV and V corresponding to the warmest air temperatures and group III containing the coldest sites in the dataset (Figs 2, 3b).

Sites across the Canadian Arctic Archipelago were separated into SOM group I, II and III. Most sampling sites from Baffin Island, Canada, were clustered in group I, while II and III groups were composed mainly of the mid- and high-arctic sites (Fig 1). All three SOM groups had the lowest mean air temperature between all regions, where mid- and high-arctic sampling sites (II and III group) had the lowest ones of the entire dataset (Table 2). Within the Canadian Arctic Archipelago, Group II had the highest number of representative taxa (Fig3a, Table 1) with *Zalutschia lingulata pauca* type (as per Medeiros and Quinlan, 2011), *Corynocera oliveri*-type, and undifferentiated *Paratanytarsus* having the highest IndVal values. The lowest number of representative taxa within Canadian Arctic Archipelago were found on Baffin Island (group I) with 4 common (IndVal > 25) and 4 rare taxa (IndVal < 25) predominantly represented by three: *Abiskomyia*, *Zalutschia* sp. C (as per Medeiros and Quinlan, 2011), and *Heterotrissocladius maeaeri*-type.2 (Fig 3a, Table 1).

Greenland sites clustered into SOM group IV were characteristic of the largest lentic systems (the highest lake area values) with intermediate air temperatures and the highest



elevations of all regions (Table 2). These ecosystems were characterized of nine common taxa, and four rare taxa (Table 1), where *Heterotrissocladius marcidus*-type, *Arctopelopia*, and *Micropsectra insignilobus*-type occurred with the highest IndVal values (Fig 3a, Table 1). Sites in Iceland clustered into group V were generally representative of the highest air temperatures in the dataset (Table 2). Iceland had the highest diversity of chironomids, with 11 common and 10 rare taxa/morphotypes. Iceland lakes were largely represented by *Heterotrissocladius grimshawi*-type, *Tanytarsus lugens*-type, and *Psectrocladius sordidellus*-type.

Examining spatial-environmental relationships with RDA obtained three significant models, one for Canada (Group I, mainly consisting of sites located on Baffin Island), Iceland, and one of two subregions within Greenland (central west Greenland). While central west Greenland was found to have one significant dbMEM (for fine-scale spatial relationship), no significant spatial relationship was found for southwest Greenland. Baffin Island and Iceland each were found to have three significant dbMEMs consisting of both fine and broad scale spatial relationships (Table 3). The low explanatory power of the RDA models is common in previous studies with similar statistical design (Heino et al., 2017) and explained as a consequence of missing environmental factors, including stochastic factors, to differentiate between sites. While spatial processes were significant for each region, variance partitioning revealed that local environmental factors were the main drivers of chironomid assemblage structure (Fig 4). Spatial processes were found to only have a small unique influence on the variation in chironomid assemblages within Baffin Island (4.3 %) and Iceland (1.9 %), while no unique variation was explained by spatial processes within either Greenlandic region (Fig 4). We note that all models had high levels of unexplained variation. Spatial processes were also found

to be significant when including all sites across the Canadian Arctic Archipelago as one single region in the VPA analysis, but this is due to the extreme spatial difference between sampling sites when considering all of northern Canada as a single spatial entity. In addition, neither the mid- or high-Arctic regions in Canada had a sufficient number of site locations paired with environmental data for analysis of spatial processes within either region using a variance partitioning approach. As such, we limit our discussion of the influence of spatial processes in the distribution of Canadian chironomid assemblages using variance partitioning to Baffin Island, which had the largest number of site locations in the Canadian dataset with the most complete data available.

## **Discussion**

As the adult form of chironomids is the main life stage for dispersal, the timing of emergence influences meta-community dynamics. Environmental controls on emergence are numerous in northern regions, predominantly governed by thermodynamic controls on lake ice, food limitation due to low productivity, wind, and air temperature (Danks et al., 2007; Medeiros and Quinlan, 2011; Dickson and Walker, 2015). Chironomid adult flight movement after emergence is predominantly less than several hundred metres from emergence origin, with some species-specific differences in dispersal distance (Ali & Fowler, 1983). So, chironomid dispersal and colonization is unlikely to be a more important control, compared to environmental conditions, on small spatial scales in water body-rich landscapes, but this may not be the case for the much longer distances between the continent and oceanic islands (Pereira et al., 2014; Ekrem et al.

2020). Although previous studies have shown chironomid assemblage structure in temperate, mid-latitude, regions is influenced by spatial processes, such as distance to source populations (Petsch et al., 2015; Árvá et al., 2015; Nicacio and Juen, 2018), temperature is a clear constraining factor for arctic communities (Oliver, 1971; Rossaro, 1991; Danks et al., 2007) and we know less about the role of biogeography there. Nonetheless, as argued in our Introduction, there are many reasons why biogeography may exert important controls on arctic chironomid distributions. For example, Allegrucci et al. (2012) showed genetically distinct chironomid populations amongst Antarctic islands, despite being geographically close, likely due to the brachypterous nature of the chironomid species, and also reflect survival in distinct glacial refugia during the past glacial maxima. Ekrem et al. (2018) also note genetically distinct chironomid populations in eastern Greenland, which supports our findings of unique assemblages in central- and south-west Greenland compared to the eastern Canadian Arctic or Iceland. The very nature of the Arctic, and its extreme environments and isolated archipelagos, may result in the evolution of behaviours that reduce dispersal, or favour species that invest fewer resources in dispersal, which would further isolate chironomid populations. Thus, biogeographic limitations could have implications for paleoecological studies seeking to use modern chironomid-environment relationships to reconstruct past environmental change. Biogeography will also contribute to shaping the major future ecosystem changes forecast for the Arctic. As arctic regions warm, ecological thresholds can be crossed where species once limited by temperature, nutrients, food availability, and light penetration through snow and ice are able to access new habitats – unless there are spatial barriers to their dispersal.

We found five distinct regions within the chironomid dataset, indicating that assemblage structure was generally different (with some exceptions) among sites on Baffin Island, the northeastern versus central Canadian Arctic Archipelago, Iceland, and central/south-west Greenland. This difference is likely attributable to biogeographic distance-related limits on dispersal, potentially highlighting an island effect of isolation (Selmi and Bouliner, 2001). Each region had several common taxa that were found in high abundances in lakes, which characterized each SOM group independently of any directed gradient analysis. Likewise, the IndVal analysis found several taxa that were low-frequency and also unique to each region. This is consistent with previous work that has documented an intriguing, island-specific patchwork of Nearctic and Palearctic species in the (overall species-impoverished) insect faunas of the North Atlantic islands. Many North American insect species are present along the west coast and in the north of Greenland, with Arctic or boreal affinities depending on latitude. European taxa increasingly predominate towards the east including on Iceland (Downes, 1988; Buckland and Dugmore, 1991; Bergersen, 1995; Panagiotakopulu, 2014). The extent to which Greenland and Iceland host endemic insect species remains debated (Bergersen, 1995; Buckland and Dugmore, 1991; Bennike, 1999; Panagiotakopulu, 2014). That being said, the diversity of Arctic chironomids based on estimates from genetic analysis of specimens is much higher than often reported (Ekrem et al. 2020), including in our analysis, as morphological identification of taxa to the species-level is difficult, and often not possible using subfossil based methods due to the lack unique morphological characteristics and localized keys.

Although biogeographic differences in assemblages are obvious between major regions in our spatially broad dataset, a VPA found that environmental filtering was the main mechanism

for shaping chironomid community structure *within* regions. This suggests that while there are numerous low-frequency taxa that are unique to each region, the overarching factor that determines chironomid species distributions within each region is a consequence of environment. For example, Delettre and Morvan (2000) showed dispersal into a terrestrial landscape rapidly declined beyond 100 m from emergence source, with very few adults captured by a distance of 500 m. Average dispersal was much further in an ‘open’ study site compared to sites with riparian vegetation (Delettre and Morvan, 2000); given the temperate focus on much chironomid dispersal research, dispersal distances may be greatly underestimated should these results be extrapolated to much more open tundra landscapes. Also, as riparian vegetation can act as a dispersal filter, as the Arctic ‘greens’ in a warming climate (Sturm et al., 2001), and taller shrub vegetation increases and becomes a dominant part of riparian zones, this may decrease adult chironomid dispersion distance.

The differentiation between sites in Iceland was most strongly related to elevation, which is not surprising as the Icelandic dataset was purposely sampled along an altitudinal gradient to capture the coldest and warmest sites in a small geographic area (Langdon et al., 2008). As such, elevation best represents the primary temperature gradient within that training set (which was not captured well by the gridded data we use for air temperatures, due to the relatively low spatial resolution of that dataset). Reflecting that problem, RDA models found that temperature (the temperature variables modeled using WorldClim) was not a defining factor for the chironomid assemblages of Iceland, despite the strong control of elevation (a proxy for temperature) there.

We found distinctions between assemblages of major biogeographic regions within the Canadian Arctic Archipelago. The SOM indicated that the Baffin region represented unique assemblages from those of the northeastern and central Arctic Archipelago, which were also distinguished by large differences in air temperature (Table 2). Summer temperatures were most significant in defining sampling sites from SOM group I (mainly Baffin Island), which had the highest temperatures of the Canadian regions, and SOM group III (mainly northeastern arctic), which has the lowest air temperatures. Elevation was also found to be important for sites in the Baffin region of Canada; however, summer temperature, lake depth, and surface area were also important in defining assemblages. It is likely that each of these factors are contributing to differences in temperatures around and within the studied lakes, consistent with the original findings of strong correlations between chironomid species assemblages and air and water temperatures at the Canadian sites (e.g., Francis et al., 2006; Medeiros and Quinlan, 2011). Temperature also presumably explains some of the spatial differentiation between Canadian assemblages and those of Iceland and central west Greenland. Canadian lakes in all three of the SOM groups are the coldest sites in the dataset and characterized by the coldest indicators known in North America, such as *Oliveridia*, *Hydrobaenus*, and *Pseudodiamesa* (Fortin et al., 2015).

Our findings are consistent with past work hinting at biogeographic effects that modify chironomid species distributions' primary response to temperature in the Arctic. Gajewski et al. (2005) found that lakes in the northern and southern parts of the Canadian Arctic Archipelago had similar assemblages compared to central islands. There were few differences in chironomid assemblages that were not attributable to temperature, and little difference associated with restrictions in dispersal that were observable. The lowest diversity occurred in the Devon and

Cornwallis Islands region of the Arctic Archipelago; however, these locations also represented the most extreme limits on available nutrients and the coldest temperatures (Gajewski et al., 2005), so environmental constraints appear just as likely to explain this low diversity as extreme northerly location. There are specific examples of taxa that have cryptic distributions, with unknown mechanisms or controls on their presence or absence from northern lakes. *Corynocera ambigua* has a well-documented ambiguous distribution (Broderson and Lindgaard, 1999; Medeiros and Quinlan, 2011). The brachypterous adult stage of this non-flying species suggests that biogeography would be an extremely limiting factor for its distribution across the arctic. Yet, surveys have found it abundant across much of the eastern Canadian Arctic, including the northern Archipelago (Fortin et al., 2015). *Abiskomyia* is strongly associated with cold temperatures but has a notably limited geographic distribution within the Canadian Arctic today (Gajewski et al., 2005; Fortin et al., 2015). Despite its abundance in Baffin Island lakes through much of the Holocene, and the presence of climatic conditions and chironomid assemblages during the last interglacial period (~125,000 years ago) that were otherwise very similar to the Holocene, *Abiskomyia* was conspicuously absent from Baffin Island during the last interglacial period (Axford et al., 2011). This demonstrates that limitations on its dispersal may be variable over time.

Of particular interest for paleoecology is how biogeographic effects on colonization might influence paleoenvironmental reconstructions based upon subfossil chironomid assemblages. For example, do biogeographic effects negate the use of geographically broad training sets to calibrate chironomid-environment relationships and to identify modern analogs to subfossil assemblages? In a major transcontinental study of modern training sets, Lotter et al.

(1999) found little fundamental difference between Holocene temperature inferences made using European versus Canadian models and datasets. While reconstructions from European lakes found better analogues for subfossil assemblages using a European calibration dataset, the Canadian calibration dataset contained more of the cold-water adapted species that were found in higher abundances during the Younger Dryas cold period ~12,000 years ago. The absence of those important cold-tolerant taxa in modern European datasets was found to be the primary factor that differentiated interpretations using models developed from Canadian versus European data-sources (Lotter et al., 1999), and argued for the utility of Canadian calibration sites despite their geographic distance and somewhat different assemblages. Likewise, in our spatially and climatically broad dataset, modern lakes in the Canadian Archipelago provide the only good analogues to late Holocene species assemblages documented in the cold mid to high Arctic regions of Greenland, where no local training sets are yet available (Axford et al., 2013; 2019) and central/south-west Greenland sites are too warm to host analogous assemblages. The training sets available from relatively warm central/south-west Greenland may provide good analogues for mid to high Arctic Greenland during past periods of warmer-than-present climate there (e.g., the mid Holocene and Last Interglacial), but are less useful than the Canadian dataset for periods colder than today. In other words, despite the fact that we find evidence for some spatial controls on chironomid species distributions across our broad study region, assemblages' best analogues may still come from faraway places with comparable temperatures. Temperature exerts a dominant control on arctic species distribution, including in our biogeographically complex study area.



## **Conclusion**

Broad-scale biogeographic effects on chironomid distributions are clearly reflected in our dataset by the distinct differences between chironomid assemblages of Iceland, central-west Greenland, and eastern Canada, defined by the presence of certain common and low-frequency, rare taxa for each region. Nonetheless, all regions also share many common taxa. The Self-Organizing Map coupled with chironomid-environment gradient analysis here show that both spatial processes and environmental controls are important within the geographically large studied area of the Canadian Arctic. Within the smaller regions of Iceland and southwest Greenland, environmental controls dominate variation in chironomid assemblages. Variance partitioning reveals that spatial processes do not influence species sorting within any of the four regions analyzed by that method (Iceland, two parts of southwest Greenland, and the Baffin region of Canada), whereas environmental parameters do – particularly temperature and variables (e.g., elevation) that are closely related to temperature. Overall, we find that environmental gradients, especially temperature, define species distributions within each region, whereas spatial processes combine with environmental gradients in determining what mix of species characterizes each broad and geographically distinct island region in our study. Paleoecological reconstructions and studies of modern ecosystem change that span across large regions of the biogeographically complex Arctic should therefore consider the roles of both biogeography and environment in controlling species distributions.

## **Acknowledgements**

This work was supported by the U.S. National Science Foundation's Office of Polar Programs (CAREER Award 1454734 to Axford) and Dalhousie University. We thank Steve Brooks and the Natural History Museum (London) for hosting a meeting on taxonomic harmonization. Melissa Chipman, Jamie McFarlin, Grace Schellinger, and Angela Self contributed to discussions on subfossil taxonomy. Data from Nuuk, Greenland, were collected as part of a Leverhulme Grant "Tipping Points: Mathematics, Metaphors and Meanings". Additional funding was received from INTERACT under the European Community's Seventh Framework Programme (Reconstructing Holocene temperature variations using chironomids from the margins of the Greenland Ice Sheet). We are also grateful for the assistance of Robert Barnett, Sally Eden and Holly Stewart in the field, and Martin West for conducting water analyses in the Durham laboratory. We thank Konrad Gajewski for providing additional original sample data used in this study, as well as David Porinchu, Sue Wilson, Erin Barley, Marie-Claude Fortin, and the many scientists and funding agencies who enabled collection of the valuable original data in the source publications used in this study. Permits and permissions required for the collection of data (where applicable) can be found in the original publications used in this study. No collection permits were required for sampling in Nuuk, Greenland.

**Data availability Statement:** The datafiles used in this study, and associated R scripts, are publically available in the corresponding authors github repository; <https://github.com/arcticecology/Greenland>. All data collected from south-west Greenland can be found in the Supplemental Information.

**Biosketch:** Dr. Medeiros research interests focus on understanding the influence of environmental stress on freshwater systems, past, present, and future. He is particularly motivated in applying new methodologies to outline the influence of environmental change on aquatic trophic systems. This includes investigation of water security through the lens of sustainability and conservation, municipal planning, and engineering for freshwater supply services. Here, the collaborative research team joins Dr. Medeiros in exploring biogeographic influences on abundance and distributions of chironomid bioindicators often used in reconstruction of past environments.

**Author contributions:** A.S.M and Y.A conceived the ideas and led the writing; D.M. analysed the data; D.R.F., E.M., I.R.W., L.H., P.L., and K.P.B. contributed to the collection of data, identification of taxa, and contributed to discussions and writing. E.M., S.W., and A.L., led the collection of data for south-west Greenland and contributed to writing, R.Q. contributed to discussions of biogeography and contributed to writing.

## Tables

**Table 1** Chironomidae taxa representative for each SOM group. Bold taxa have IndVal values of more than 25% ( $p < 0.01$ ).

Group	Taxa	IndVal	$p$	Group	Taxa	IndVal	$p$
1	<b>Abiskomyia</b>	<b>69.7</b>	<b>0.0002</b>	3	<b>Pseudodiamesa</b>	<b>74.9</b>	<b>0.0002</b>
1	<b>Zalutschia sp. C</b>	<b>57.5</b>	<b>0.0002</b>	3	<b>Oliveridia</b>	<b>67.4</b>	<b>0.0002</b>
1	<b>Heterotrissocladius maeeri-type 2</b>	<b>54.2</b>	<b>0.0002</b>	3	<b>Micropsectra radialis-type</b>	<b>64.5</b>	<b>0.0002</b>
1	<b>Mesocricotopus</b>	<b>47.3</b>	<b>0.0002</b>	3	<b>Paracladius</b>	<b>26.8</b>	<b>0.0018</b>
1	<b>Heterotrissocladius maeeri-type 1</b>	<b>45.9</b>	<b>0.0002</b>	3	<i>Orthocladius type S</i>	21.9	0.0026
1	<b>Protanypus</b>	<b>33.7</b>	<b>0.0002</b>	3	<i>Chaetocladius</i>	8.4	0.0566
1	<i>Parakiefferiella nigra-type</i>	23.1	0.0004	4	<b>Heterotrissocladius marcidus-type</b>	<b>63.4</b>	<b>0.0002</b>
1	<i>Heterotrissocladius subpilosus-type</i>	19.1	0.0088	4	<b>Arctopelopia</b>	<b>52</b>	<b>0.0002</b>
1	<i>Tanytarsus pallidicornis-type</i>	15.9	0.0078	4	<b>Procladius</b>	<b>45.1</b>	<b>0.0002</b>
1	<i>Psectrocladius psilopterus-type</i>	9.7	0.0086	4	<b>Dicrotendipes</b>	<b>39.2</b>	<b>0.0002</b>
1	<i>Tvetenia</i>	7.4	0.0524	4	<b>Chironomus anthracinus-type</b>	<b>33.9</b>	<b>0.0002</b>
1	<i>Corynoneura edwardsi-type</i>	4.8	0.0456	4	<b>Cricotopus type P</b>	<b>33.3</b>	<b>0.0002</b>
1	<i>Diamesa aberrata-type</i>	4.8	0.0484	4	<b>Micropsectra insignilobus-type</b>	<b>33.2</b>	<b>0.0002</b>
1	<i>Psectrocladius (Allopsectrocladius)</i>	4.8	0.0456	4	<b>Micropsectra contracta-type</b>	<b>32.8</b>	<b>0.0002</b>
1	<i>Tanytarsus lactescens-type</i>	4.8	0.0492	4	<b>Ablabesmyia</b>	<b>32.3</b>	<b>0.0002</b>
2	<b>Zalutschia lingulata pauca type</b>	<b>59.4</b>	<b>0.0002</b>	4	<i>Tanytarsus mendax-type</i>	20.3	0.0006
2	<b>Corynocera oliveri-type</b>	<b>56</b>	<b>0.0002</b>	4	<i>Psectrocladius septentrionalis-type</i>	12.4	0.007
2	<b>Paratanytarsus</b>	<b>39.9</b>	<b>0.0002</b>	5	<b>Heterotrissocladius</b>	<b>87.1</b>	<b>0.0002</b>

	undifferentiated				<i>grimshawi</i> -type		
2	<i>Paratanytarsus penicillatus</i> -type	39.7	0.0002	5	<i>Tanytarsus lugens</i> -type	82.1	0.0002
2	<i>Tanytarsus lugens</i> -type	36	0.0002	5	<i>Psectrocladius sordidellus</i> -type	49.8	0.0002
2	<i>Corynocera ambigua</i>	31.6	0.0002	5	<i>Thienemanniella</i>	45.8	0.0002
2	<i>Sergentia coracina</i> -type	30.4	0.0006	5	<i>Cricotopus tremulus</i> group	40.6	0.0002
2	<i>Limnophyes</i> / <i>Paralimnophyes</i>	28.9	0.0006	5	<i>Cricotopus sylvestris</i> -type	37.7	0.0002
2	<i>Corynoneura arctica</i> -type	26.9	0.0024	5	<i>Eukiefferiella fittkau</i> -type	34.1	0.0002
2	<i>Stictochironomus</i>	26.2	0.0002	5	<i>Diamesa zernyi</i> -type	32.5	0.0002
2	<i>Cricotopus intersectus</i> -type	23.7	0.0014	5	<i>Eukiefferiella</i>	30.9	0.0002
2	<i>Orthocladius consobrinus</i> -type	18.8	0.0138	5	<i>Paracladopelma</i>	24.8	0.0012
2	<i>Orthocladius</i> type I	18.5	0.0112	5	<i>Cricotopus cylindraceus</i> -type	23.6	0.002
2	<i>Hydrobaenus</i>	18.2	0.0004	5	<i>Psectrocladius barbimanus</i> -type	23.2	0.0018
2	<i>Metriocnemus</i>	15	0.002	5	<i>Orthocladius oliveri</i> -type	21.2	0.0008
2	<i>Cricotopus laricomalis</i> -type	12.4	0.0116	5	<i>Rheocricotopus</i>	17	0.0006
2	<i>Parakiefferiella triquetra</i> -type	10.6	0.0384	5	<i>Paraphenocladius</i>	16	0.0014
2	<i>Cladotanytarsus mancus</i> -type	9.9	0.0036	5	<i>Pseudosmittia</i>	15.1	0.0034
2	<i>Parakiefferiella</i> type B	9	0.029	5	<i>Diamesa arctica</i> -type	11.3	0.0046
2	<i>Psectrocladius calcaratus</i> -type	7.9	0.0158	5	<i>Parochlus</i>	11.3	0.0048

**Table 2.** Mean values of environmental parameters ( $\pm 1$  standard error) for each SOM group. Values within the same row not sharing a common superscript letter are significantly different: <sup>a,b,c</sup>  $p < 0.01$ . If superscript letters are missing, all the values are significantly different. Environmental patterns classified by SOM over different groups were confirmed by non-parametric ANOVA (Kruskal–Wallis tests) and a Mann–Whitney post-hoc test for pair-wise comparisons. Environmental parameters are as follows; Elevation (Elev), Surface area (Area), Conductivity (Cond), Depth of sample (Depth), Annual precipitation (Prec), Mean annual temperature (MEAN), and Mean temperature of the warmest quarter (SumT).

Env.	Group I	Group II	Group III	Group IV	Group V
<b>Elev</b>	213.43 $\pm$	199.84 $\pm$	212.79 $\pm$	314.84 $\pm$ 23.92	221.39 $\pm$
(m a.s.l.)	20.26 <sup>a</sup>	21.91 <sup>ab</sup>	32.10 <sup>a</sup>		24.17 <sup>a</sup>
<b>Area (km<sup>2</sup>)</b>	0.40 $\pm$	0.13 $\pm$ 0.54	0.65 $\pm$ 0.27 <sup>a</sup>	23.216 $\pm$ 4.74	0.45 $\pm$ 0.11 <sup>a</sup>
	0.99 <sup>ab</sup>				
<b>Cond</b>	64.41 $\pm$	118.77 $\pm$	116.84 $\pm$	329.73 $\pm$	102.58 $\pm$
( $\mu$ S/cm)	18.32a	33.04ab	27.30b	106.93b	6.96b
<b>Depth</b>	10.70 $\pm$	3.24 $\pm$ 0.39 <sup>c</sup>	8.02 $\pm$ 0.97 <sup>a</sup>	12.82 $\pm$ 1.44 <sup>b</sup>	6.51 $\pm$ 1.57 <sup>c</sup>
(m)	1.05 <sup>ab</sup>				
<b>PREC</b>	253.97 $\pm$	147.84 $\pm$	128.17 $\pm$	429.13 $\pm$ 20.46	915.32 $\pm$
(mm year <sup>-1</sup> )	15.45	16.97 <sup>a</sup>	6.93 <sup>a</sup>		16.80
<b>MEAN</b>	-13.42 $\pm$	-16.96 $\pm$ 0.61 <sup>a</sup>	-17.73 $\pm$ 0.31 <sup>a</sup>	-4.40 $\pm$ 0.27	2.18 $\pm$ 0.14
(°C)	0.42				
<b>SumT</b>	3.14 $\pm$ 0.14 <sup>a</sup>	2.79 $\pm$ 0.22 <sup>a</sup>	1.83 $\pm$ 0.14	6.62 $\pm$ 0.13	8.02 $\pm$ 0.12
(°C)					

**Table 3** Significant environmental parameters and the distance-based Moran's eigenvector maps (dbMEMs) as spatial variables in RDA models, selected using forward selection, for the four studied regions. Low eigenvalues of dbMEMs indicate fine-scale relationships between neighbouring sampling sites, while high eigenvalues of dbMEMs indicate coarse, broad-scale, patterns of relationships between studied sites.

<b>Canada (Baffin)</b>	<b>Iceland</b>	<b>Central-west Greenland (Søndre Strømfjord)</b>	<b>South-west Greenland (Nuuk)</b>
Elevation	Elevation	Depth	Depth
Lake area	Lake area	Mean temperature	
Depth	Precipitation	Conductivity	
Mean temperature	TN	TN	
	TC		
MEM 1,2,6	MEM 1,2,4	MEM1	

## Figures

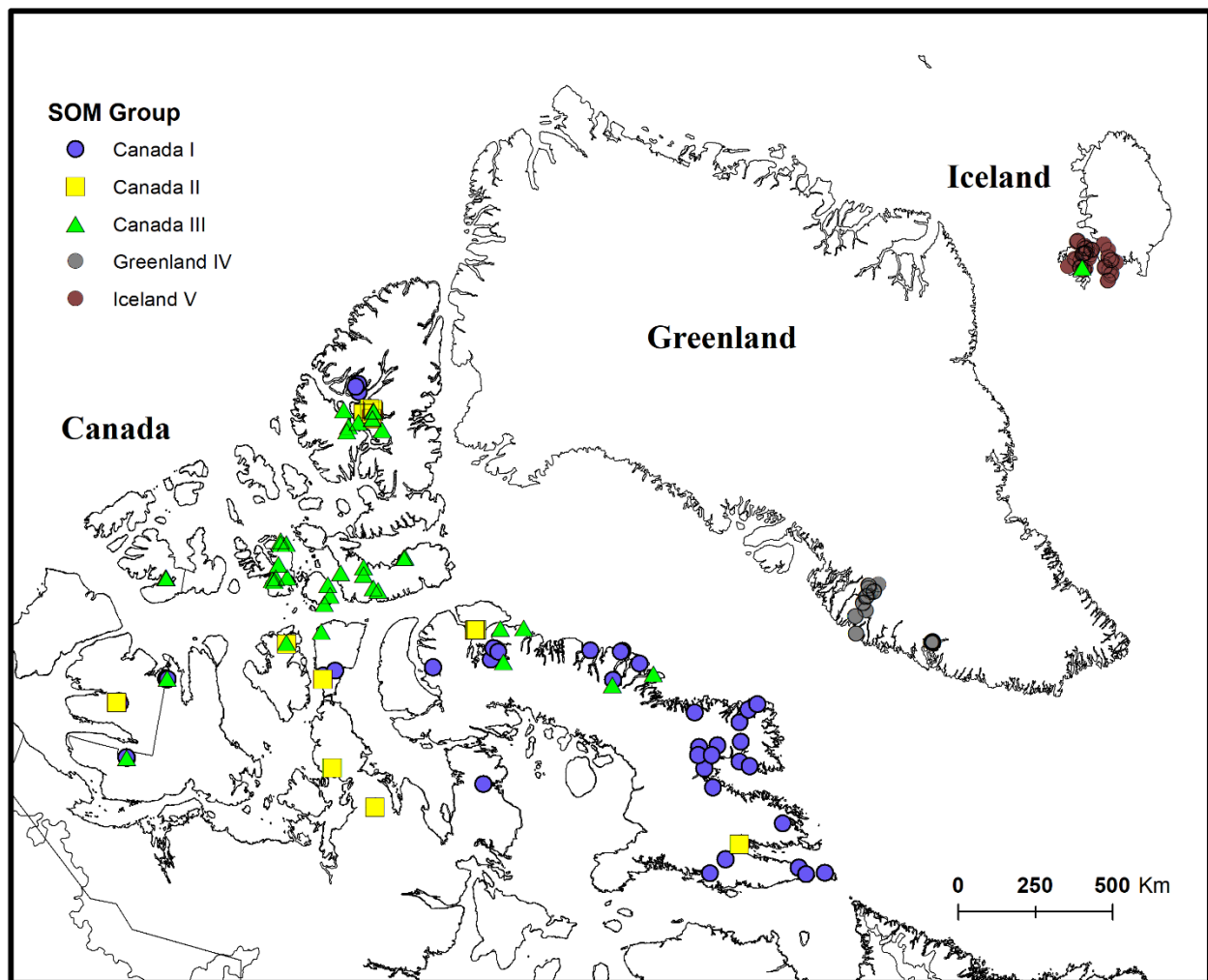


Fig. 1 Sampling sites in Canada, southwest Greenland, and northwest Iceland. Colored symbols indicate Self Organizing Map (SOM) group classification.



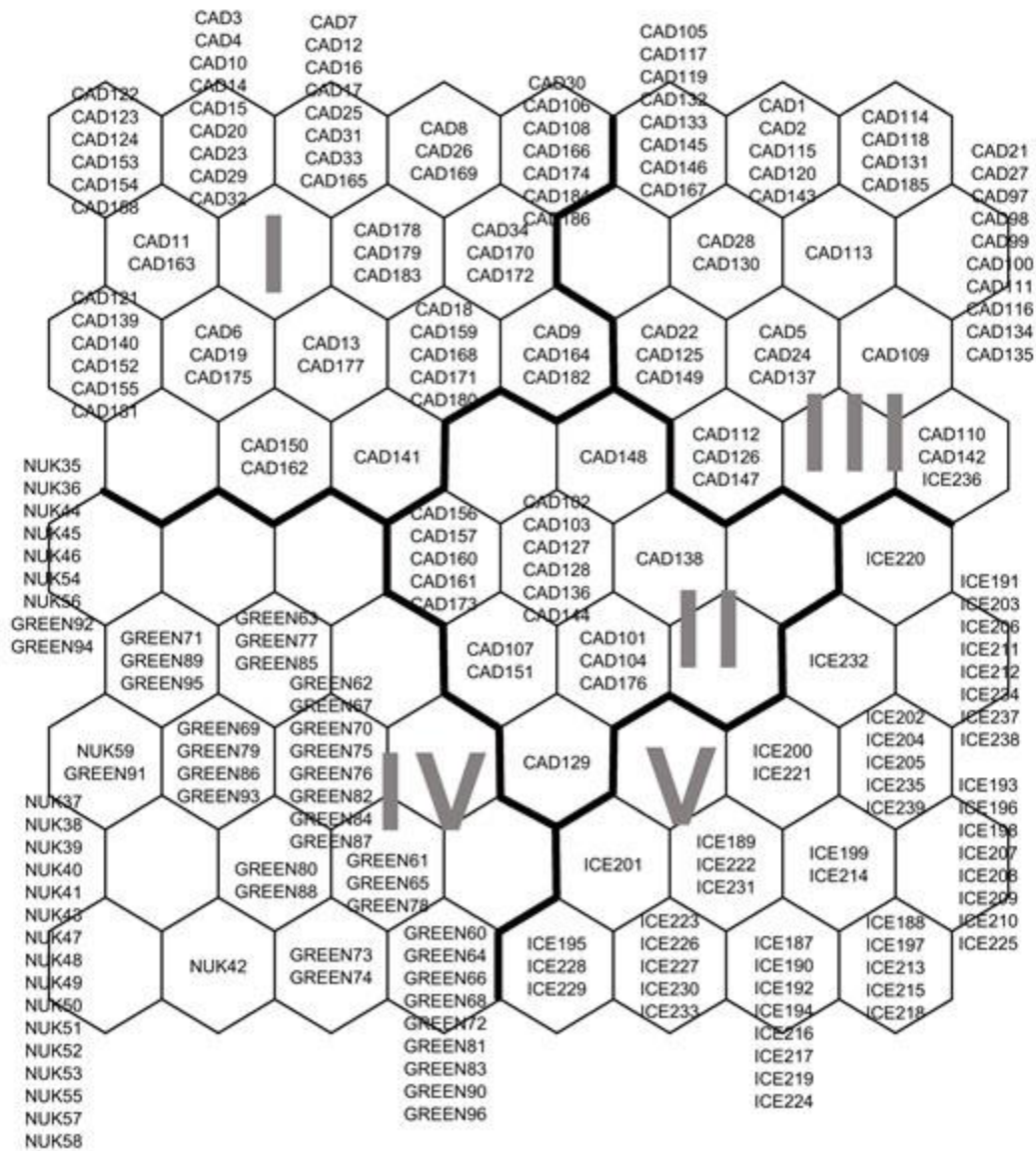
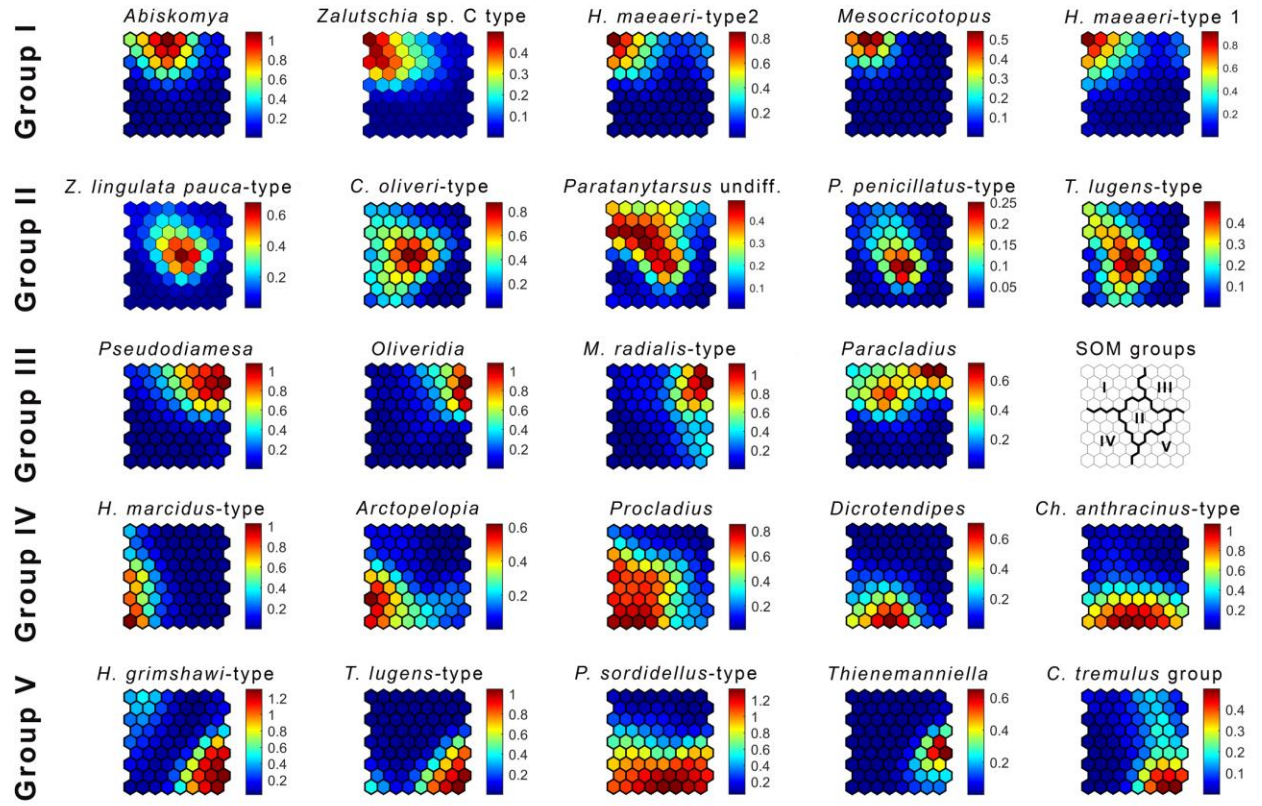


Fig. 2 Self-Organizing Map (SOM), with classification pattern based on chironomid assemblages. Latin numbers indicate the subset of neurons obtained by k-means method. The labels assigned to each neuron stand for particular sampling sites where first capital letters indicate the region (CAD-Baffin Island in Canada, ICE -, northwest Iceland, and GREEN / NUK- Søndre Strømfjord in central-west Greenland, NUK – Nuuk region in south-west Greenland).

A)



B)

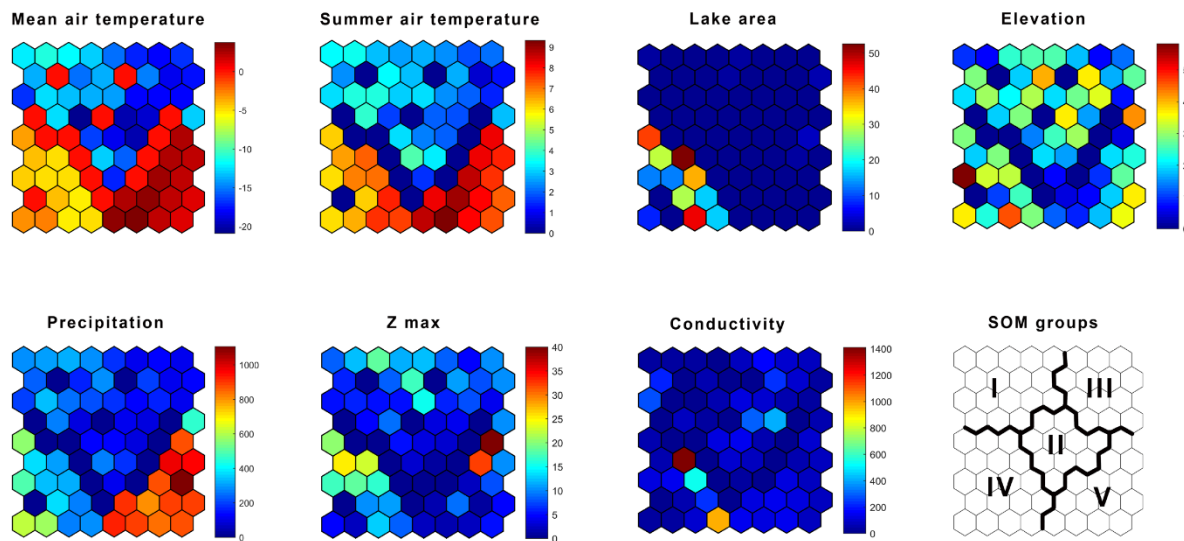


Fig. 3 Component planes presenting the distributional pattern of a) indicator taxa and b) environmental parameters (July- Air temperature in July, MEAN- mean air temperature, SumT- Summer air temperature, Water-water temperature, Depth-Lake depth, Elev-Elevation) over the trained SOM. Colour gradient indicates a) relative abundance of indicator taxa or b) values of environmental parameters, with red and blue for highest and lowest values, respectively. Units without assigned samples (empty neurons) are also coloured in blue.

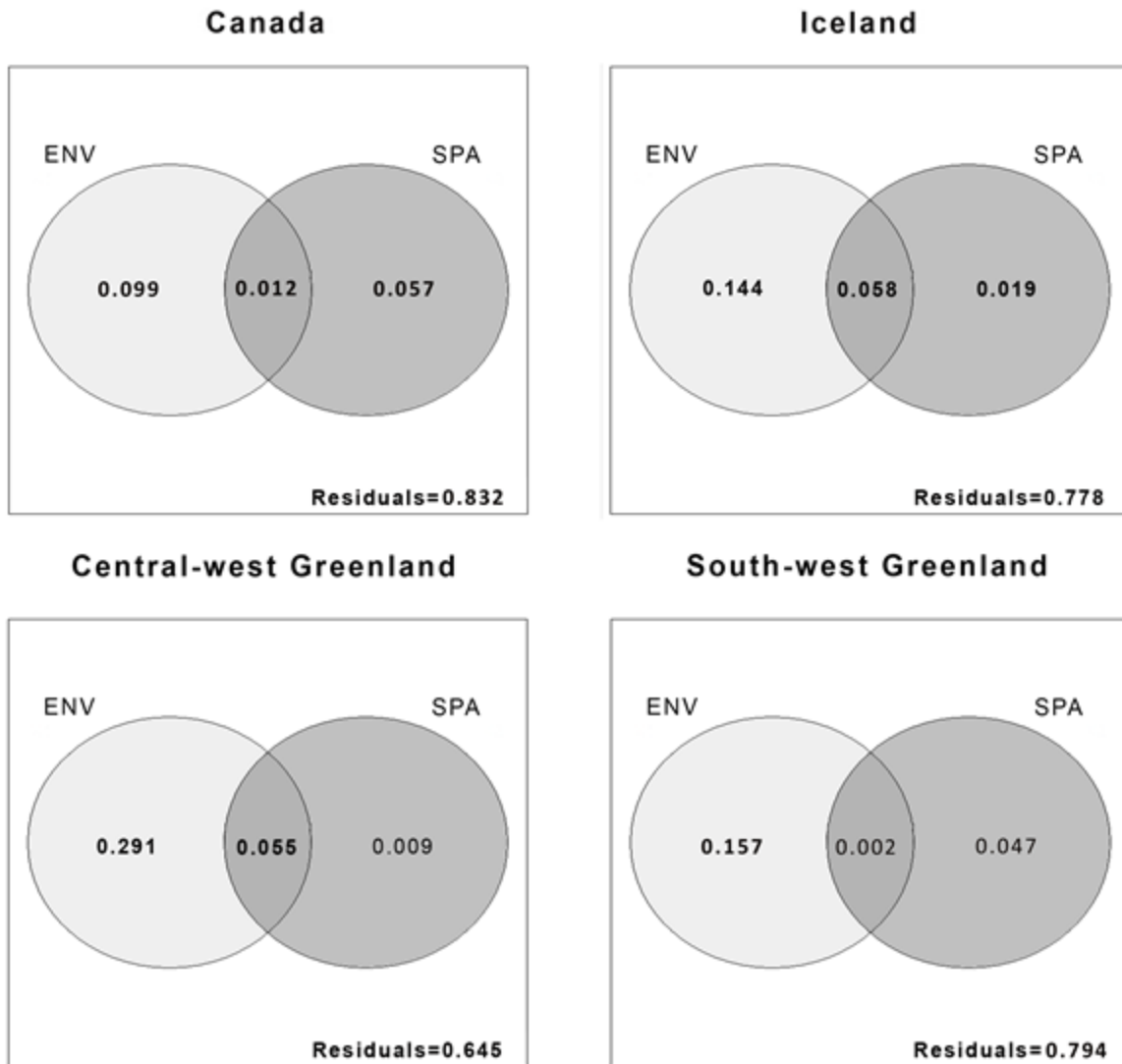


Fig. 4 Variation partitioning for chironomid meta-community structure from four regions; Baffin Island in Canada, northwest Iceland, south-west Greenland, and central-west Greenland explained by environmental parameters (X1) and spatial variables (X2). Residuals are presented in the lower right corner of each Venn diagram.

## References

- Ali, A., & Fowler, R. C. (1983). Prevalence and dispersal of pestiferous Chironomidae in a lake front city of central Florida. *Mosquito News*, 43(1), 55-59.
- Allegrucci, G., Carchini, G., Convey, P., & Sbordoni, V. (2012). Evolutionary geographic relationships among orthocladine chironomid midges from maritime Antarctic and sub-Antarctic islands. *Biological Journal of the Linnean Society*, 106(2), 258-274.
- Árva, D., Tóth, M., Horváth, H., Nagy, S. A., & Specziár, A. (2015). The relative importance of spatial and environmental processes in distribution of benthic chironomid larvae within a large and shallow lake. *Hydrobiologia*, 742(1), 249-266.
- Axford, Y., Briner, J. P., Francis, D. R., Miller, G. H., Walker, I. R., & Wolfe, A. P. (2011). Chironomids record terrestrial temperature changes throughout arctic interglacials of the past 200,000 yr. *GSA Bulletin*, 123(7-8), 1275-1287.
- Axford, Y., Losee, S., Briner, J. P., Francis, D. R., Langdon, P. G., & Walker, I. R. (2013). Holocene temperature history at the western Greenland Ice Sheet margin reconstructed from lake sediments. *Quaternary Science Reviews*, 59(1), 87-100.
- Axford, Y., Lasher, G. E., Kelly, M. A., Osterberg, E. C., Landis, J., Schellinger, G. C., ... & Francis, D. R. (2019). Holocene temperature history of northwest Greenland—With new ice cap constraints and chironomid assemblages from Deltasø. *Quaternary Science Reviews*, 215(1), 160-172.

- Balasubramaniam, A. M., Medeiros, A. S., Turner, K. W., Hall, R. I., & Wolfe, B. B. (2017). Biotic responses to multiple aquatic and terrestrial gradients in shallow subarctic lakes (Old Crow Flats, Yukon, Canada). *Arctic Science*, 3(2), 277-300.
- Beattie, D. M. (1981). Investigations into the occurrence of midge plagues in the vicinity of Harderwijk (Netherlands). *Hydrobiologia*, 80(2), 147-159.
- Bennike, O. (1999). Colonisation of Greenland by plants and animals after the last ice age: a review. *Polar Record*, 35(195), 323-336.
- Bergersen, R. (1995). Special Paper: Is Greenland a Zoogeographical Unit of Its Own? *Journal of Biogeography*, 22(1), 1-6.
- Bitušík, P., Svitok, M., Novikmec, M., Trnková, K. & Hamerlík, L. 2017. A unique way of passive dispersal of aquatic invertebrates by wind: Chironomid larvae are traveling in fragments of aquatic mosses. *Limnologica*, 63(1), 119-121.
- Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Forward selection of explanatory variables. *Ecology*, 89(9), 2623-2632.
- Borcard, D., & Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological modelling*, 153(1-2), 51-68.
- Borcard, D., Legendre, P., Avois-Jacquet, C., & Tuomisto, H. (2004). Dissecting the spatial structure of ecological data at multiple scales. *Ecology*, 85(7), 1826-1832.

- Brochmann, C., Gabrielsen, T. M., Nordal, I., Landvik, J. Y., & Elven, R. (2003). Glacial survival or tabula rasa? The history of North Atlantic biota revisited. *Taxon*, 52(3), 417-450.
- Brodersen, K. P., & Lindegaard, C. (1999). Classification, assessment and trophic reconstruction of Danish lakes using chironomids. *Freshwater Biology*, 42(1), 143-157.
- Brodersen, K.P. & Anderson, N.J., 2002. Distribution of chironomids (Diptera) in low arctic West Greenland lakes: trophic conditions, temperature and environmental reconstruction. *Freshwater Biology*, 47(6), 1137-1157.
- Brooks, S. J., & Birks, H. J. B. (2001). Chironomid-inferred air temperatures from Lateglacial and Holocene sites in north-west Europe: progress and problems. *Quaternary Science Reviews*, 20(16-17), 1723-1741.
- Brooks, S. J., Langdon, P. G., & Heiri, O. (2007). Using and identifying chironomid larvae in palaeoecology, QRA Technical Guide No 10. *Quaternary Research Association, London*.
- Buckland, P., & Dugmore, A. (1991). 'If this is a refugium, why are my feet so bloody cold?' The origins of the Icelandic biota in the light of recent research. In *Environmental change in Iceland: past and present* (pp. 107-125). Springer, Dordrecht.
- Chon, T. S., Park, Y. S., Moon, K. H., & Cha, E. Y. (1996). Patternizing communities by using an artificial neural network. *Ecological modelling*, 90(1), 69-78.
- Clarke, K. R., & Warwick, R. M. (1994). An approach to statistical analysis and interpretation. *Change in Marine Communities*, 2(1), 117-143.

- Clesceri, L. S., Greenberg, A. E., & Eaton, A. D. 1998. *Standard methods for the examination of water and wastewater (20th ed.)*: Washington, DC, American Public Health Association.
- Crump, S. E., Miller, G. H., Power, M., Sepúlveda, J., Dildar, N., Coghlan, M., & Bunce, M. (2019). Arctic shrub colonization lagged peak postglacial warmth: Molecular evidence in lake sediment from Arctic Canada. *Global Change Biology*, 25(12), 4244-4256.
- Danks, H. V. (2007). The elements of seasonal adaptations in insects. *The Canadian Entomologist*, 139(1), 1-44.
- Dean, W.E., 1974. Determination of carbonate and organic matter in calcareous sediments and sedimentary rocks by loss on ignition; comparison with other methods. *Journal of Sedimentary Research*, 44(1), 242-248.
- Delettre, Y. R. (1988). Chironomid wing length, dispersal ability and habitat predictability. *Ecography*, 11(3), 166-170.
- Dickson, TR, IR Walker (2015) Midge (Diptera: Chironomidae and Ceratopogonidae) emergence responses to temperature: experiments to assess midges' capacity as paleotemperature indicators. *Journal of Paleolimnology*, 53(2), 165-176
- Delettre, Y. R., & Morvan, N. (2000). Dispersal of adult aquatic Chironomidae (Diptera) in agricultural landscapes. *Freshwater Biology*, 44(3), 399-411.
- Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological monographs*, 67(3), 345-366.



- Dickson, T. R., Bos, D. G., Pellatt, M. G., & Walker, I. R. (2014). A midge-salinity transfer function for inferring sea level change and landscape evolution in the Hudson Bay Lowlands, Manitoba, Canada. *Journal of Paleolimnology*, 51(3), 325-341.
- Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, 196(3-4), 483-493.
- Dray, S., Legendre, P., & Blanchet, G. (2009). packfor: Forward Selection with permutation (Canoco p. 46). R package version 0.0-8.
- Dray, S., Péliissier, R., Couteron, P., Fortin, M. J., Legendre, P., Peres-Neto, P. R., ... & Dufour, A. B. (2012). Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs*, 82(3), 257-275
- Eggermont, H., & Heiri, O. (2012). The chironomid-temperature relationship: expression in nature and palaeoenvironmental implications. *Biological Reviews*, 87(2), 430-456.
- Engels, S., Medeiros, A. S., Axford, Y., Brooks, S. J., Heiri, O., Luoto, T. P., ... & Self, A. E. (2019). Temperature change as a driver of spatial patterns and long-term trends in chironomid (Insecta: Diptera) diversity. *Global Change Biology*, 26(3), 1155-1169.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302-4315.

- Fortin, M. C., Medeiros, A. S., Gajewski, K., Barley, E. M., Larocque-Tobler, I., Porinchu, D. F., & Wilson, S. E. (2015). Chironomid-environment relations in northern North America. *Journal of Paleolimnology*, 54(2-3), 223-237.
- Francis, D. R., Wolfe, A. P., Walker, I. R., & Miller, G. H. (2006). Interglacial and Holocene temperature reconstructions based on midge remains in sediments of two lakes from Baffin Island, Nunavut, Arctic Canada. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 236(1-2), 107-124.
- Frisch, D., Green, A. J., & Figuerola, J. (2007). High dispersal capacity of a broad spectrum of aquatic invertebrates via waterbirds. *Aquatic Sciences*, 69(4), 568-574.
- Gajewski, K., Bouchard, G., Wilson, S. E., Kurek, J., & Cwynar, L. C. (2005). Distribution of Chironomidae (Insecta: Diptera) head capsules in recent sediments of Canadian Arctic lakes. *Hydrobiologia*, 549(1), 131-143.
- Heino, J., Soininen, J., Alahuhta, J., Lappalainen, J., & Virtanen, R. (2017). Metacommunity ecology meets biogeography: effects of geographical region, spatial dynamics and environmental filtering on community structure in aquatic organisms. *Oecologia*, 183(1), 121-137.
- Heiri, O., Lotter, A.F., Lemcke, G., (2001). Loss on ignition as a method for estimating organic and carbonate content in sediments: reproducibility and comparability of results. *Journal of Paleolimnology*, 25(1), 101-110.
- Jain, A. K., & Dubes, R. C. (1988). *Algorithms for clustering data*. Prentice-Hall, Inc..

- Johansen, S., & Hytteborn, H. (2001). A contribution to the discussion of biota dispersal with drift ice and driftwood in the North Atlantic. *Journal of Biogeography*, 28(1), 105-115.
- Kaufman, D. S., Axford, Y. L., Henderson, A. C., McKay, N. P., Oswald, W. W., Saenger, C., ... & Hu, F. S. (2016). Holocene climate changes in eastern Beringia (NW North America)—A systematic review of multi-proxy evidence. *Quaternary Science Reviews*, 147(1), 312-339.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1), 59-69.
- Langdon, P. G., Ruiz, Z. O. E., Brodersen, K. P., & Foster, I. D. (2006). Assessing lake eutrophication using chironomids: understanding the nature of community response in different lake types. *Freshwater Biology*, 51(3), 562-577.
- Langdon, P. G., Holmes, N., & Caseldine, C. J. (2008). Environmental controls on modern chironomid faunas from NW Iceland and implications for reconstructing climate change. *Journal of Paleolimnology*, 40(1), 273-293.
- Legendre P, Legendre L (1998) Numerical ecology, 2nd English edn. Elsevier, Amsterdam
- Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129(2), 271-280.
- Lotter, A. F., Walker, I. R., Brooks, S. J., & Hofmann, W. (1999). An intercontinental comparison of chironomid palaeotemperature inference models: Europe vs North America. *Quaternary Science Reviews*, 18(6), 717-735.

- Luoto, T. P., & Ojala, A. E. (2018). Controls of climate, catchment erosion and biological production on long-term community and functional changes of chironomids in high arctic lakes (Svalbard). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 505(1), 63-72.
- McCune, B. and M. J. Mefford. (2011). PC-ORD. Multivariate Analysis of Ecological Data. Version 6.08 MjM Software, Gleneden Beach, Oregon, U.S.A.
- Medeiros, A. S., & Quinlan, R. (2011). The distribution of the Chironomidae (Insecta: Diptera) along multiple environmental gradients in lakes and ponds of the eastern Canadian Arctic. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(9), 1511-1527.
- Medeiros, A. S., Friel, C. E., Finkelstein, S. A., & Quinlan, R. (2012). A high resolution multi-proxy record of pronounced recent environmental change at Baker Lake, Nunavut. *Journal of Paleolimnology*, 47(4), 661-676.
- Milošević, D., Simić, V., Stojković, M., Čerba, D., Mančev, D., Petrović, A., & Paunović, M. (2013). Spatio-temporal pattern of the Chironomidae community: toward the use of non-biting midges in bioassessment programs. *Aquatic Ecology*, 47(1), 37-55.
- Murray, D. A., & O'Connor, J. P. (1992). Observations on the distribution of some chironomids (Diptera) and caddisflies (Trichoptera) in the Killarney area, Ireland. *Netherland Journal of Aquatic Ecology*, 26(2-4), 393-396.
- Namayandeh, A., & Quinlan, R. (2011). Benthic macroinvertebrate communities in Arctic lakes and ponds of central Nunavut, Canada. *Arctic, Antarctic, and Alpine Research*, 43(3), 417-428.

- Nazarova, L., Self, A. E., Brooks, S. J., van Hardenbroek, M., Herzsuh, U., & Diekmann, B. (2015). Northern Russian chironomid-based modern summer temperature data set and inference models. *Global and Planetary Change*, 134(1), 10-25.
- Nicacio, G., & Juen, L. (2018). Relative roles of environmental and spatial constraints in assemblages of Chironomidae (Diptera) in Amazonian floodplain streams. *Hydrobiologia*, 820(1), 201-213.
- Novichkova, A., Chertoprud, E., & Gíslason, G. M. (2014). Freshwater Crustacea (Cladocera, Copepoda) of Iceland: taxonomy, ecology, and biogeography. *Polar Biology*, 37(12), 1755-1767.
- Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., & Oksanen, M. J. (2007). The Vegan Package—Community Ecology Package. R package version 2.5-6.
- Oliver, D. R. (1971). Life history of the Chironomidae. *Annual Review of Entomology*, 16(1), 211-230.
- Panagiotakopulu, E. (2014). Hitchhiking across the North Atlantic—Insect immigrants, origins, introductions and extinctions. *Quaternary International*, 341(1), 59-68.
- Park, Y. S., Céréghino, R., Compin, A., & Lek, S. (2003). Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters. *Ecological Modelling*, 160(3), 265-280.

- Park, Y. S., Chon, T. S., Kwak, I. S., & Lek, S. (2004). Hierarchical community classification and assessment of aquatic ecosystems using artificial neural networks. *Science of the total Environment*, 327(1-3), 105-122.
- Penczak, T., Kruk, A., & Galicka, W. (2012). Implementation of a self-organizing map for investigation of impoundment impact on fish assemblages in a large, lowland river: Long-term study. *Ecological Modelling*, 227(1), 64-71.
- Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, 87(10), 2614-2625.
- Petsch, D. K., Pinha, G. D., Dias, J. D., & Takeda, A. M. (2015). Temporal nestedness in Chironomidae and the importance of environmental and spatial factors in species rarity. *Hydrobiologia*, 745(1), 181-193.
- Primack, R. B., Miller-Rushing, A. J., Corlett, R. T., Devictor, V., Johns, D., Loyola, R., ... & Pejchar, L. (2018). Biodiversity gains? The debate on changes in local-vs global-scale species richness. *Biological Conservation*, 219(1), A1-A3.
- Quinlan, R., Douglas, M. S., & Smol, J. P. (2005). Food web changes in arctic ecosystems related to climate warming. *Global Change Biology*, 11(8), 1381-1386.
- Selmi, S. D., & Boulinier, T. (2001). Ecological biogeography of Southern Ocean islands: the importance of considering spatial issues. *The American Naturalist*, 158(4), 426-437.
- Sturm, M., Racine, C. & Tape, K. (2001). Increasing shrub abundance in the Arctic. *Nature*, 411(6837), 546-547.

- Rossaro, B. (1991). Chironomids and water temperature. *Aquatic Insects*, 13(2), 87-98.
- Rundgren, M., & Ingólfsson, O. (1999). Plant survival in Iceland during periods of glaciation? *Journal of Biogeography*, 26(2), 387-396.
- Taylor, D. J., Ballinger, M. J., Medeiros, A. S., & Kotov, A. A. (2016). Climate-associated tundra thaw pond formation and range expansion of boreal zooplankton predators. *Ecography*, 39(1), 43-53.
- Vellend, M., Baeten, L., Myers-Smith, I. H., Elmendorf, S. C., Beauséjour, R., Brown, C. D., ... & Wipf, S. (2013). Global meta-analysis reveals no net change in local-scale plant biodiversity over time. *Proceedings of the National Academy of Sciences*, 110(48), 19456-19459.
- Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, 11(3), 586-600.
- Walker, I. R., Smol, J. P., Engstrom, D. R., & Birks, H. J. B. (1991). An assessment of Chironomidae as quantitative indicators of past climatic change. *Canadian Journal of Fisheries and Aquatic Sciences*, 48(6), 975-987.
- Westover, K. S., Moser, K. A., Porinchu, D. F., MacDonald, G. M., & Wang, X. (2009). Physical and chemical limnology of a 61-lake transect across mainland Nunavut and southeastern Victoria Island, Central Canadian Arctic. *Fundamental and Applied Limnology/Archiv für Hydrobiologie*, 175(2), 93-112.

Wright, H. E., 1990. An improved Hongve sampler for surface sediments. *Journal of Paleolimnology*, 4(1), 91-92.